



Lawlor, D., Lewcock, M., Jones, L., Rollings, C., Yip, V., Smith, D., Pearson, R., Johnson, L., Millard, L., Patel, N., Skinner, A., & Tilling, K. (2019). The second generation of The Avon Longitudinal Study of Parents and Children (ALSPAC-G2): a cohort profile. *Wellcome Open Research*, 4(36). <https://doi.org/10.12688/wellcomeopenres.15087.1>

Publisher's PDF, also known as Version of record

License (if available):
CC BY

Link to published version (if available):
[10.12688/wellcomeopenres.15087.1](https://doi.org/10.12688/wellcomeopenres.15087.1)

[Link to publication record in Explore Bristol Research](#)
PDF-document

This is the final published version of the article (version of record). It first appeared online via Wellcome at <https://wellcomeopenresearch.org/articles/4-36/v1>. Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>



RESEARCH ARTICLE

The second generation of The Avon Longitudinal Study of Parents and Children (ALSPAC-G2): a cohort profile [version 1; referees: awaiting peer review]

Deborah A. Lawlor¹⁻³, Melanie Lewcock^{2,4}, Louise Rena-Jones^{2,4}, Claire Rollings^{2,4}, Vikki Yip^{2,4}, Daniel Smith^{2,4}, Rebecca M. Pearson^{1-3,5}, Laura Johnson^{1,3,6}, Louise A. C. Millard^{1,2,7}, Nashita Patel⁸, Andy Skinner^{1,2}, Kate Tilling¹⁻³, ALSPAC Executive

¹MRC Integrative Epidemiology Unit at the University of Bristol, University of Bristol, Bristol, BS8 2BN, UK

²Population Health Science, Bristol Medical School, University of Bristol, Bristol, UK

³NIHR Bristol Biomedical Research Centre, Bristol, UK

⁴ALSPAC, University of Bristol, Bristol, UK

⁵Centre for Academic Mental Health, University of Bristol, Bristol, UK

⁶Centre for Exercise, Nutrition and Health Science, School for Policy Studies, University of Bristol, Bristol, UK

⁷Intelligent Systems Laboratory, University of Bristol, Bristol, UK

⁸Department of Women and Children's Health, School of Life Course Sciences, Kings College London, London, UK

v1 First published: 20 Feb 2019, 4:36 (<https://doi.org/10.12688/wellcomeopenres.15087.1>)

Latest published: 20 Feb 2019, 4:36 (<https://doi.org/10.12688/wellcomeopenres.15087.1>)

Abstract

Background: The Avon Longitudinal Study of Parents and Children-Generation 2 (ALSPAC-G2) was set up to provide a unique multi-generational cohort. It builds on the existing ALSPAC resource, which recruited 14,541 pregnancies to women resident in the South West of England who were expected to deliver between 01/04/1991 and 31/12/1992. Those women and their partners (Generation 0; ALSPAC-G0) and their offspring (ALSPAC-G1) have been followed for the last 26 years. This profile describes recruitment and data collection on the next generation (ALSPAC-G2)—the grandchildren of ALSPAC-G0 and children of ALSPAC-G1.

Recruitment: Recruitment began on the 6th of June 2012 and we present details of recruitment and participants up to 30th June 2018 (~6 years). We knew at the start of recruitment that some ALSPAC-G1 participants had already become parents and ALSPAC-G2 is an open cohort; we recruit at any age. We hope to continue recruiting until all ALSPAC-G1 participants have completed their families. Up to 30th June 2018 we recruited 810 ALSPAC-G2 participants from 548 families. Of these 810, 389 (48%) were recruited during their mother's pregnancy, 287 (35%) before age 3 years, 104 (13%) between 3-6 years and 30 (4%) after 6 years. Over 70% of those invited to early pregnancy, late pregnancy, second week of life, 6-, 12- and 24-month assessments (whether for their recruitment, or a follow-up, visit) have attended, with attendance being over 60% for subsequent visits up to 7 years (to few are eligible for the 9- and 11-year assessments to analyse).

Open Peer Review

Referee Status: Awaiting Peer

Review

Any reports and responses or comments on the article can be found at the end of the article.

Data collection: We collect a wide-range of social, lifestyle, clinical, anthropometric and biological data on all family members repeatedly. Biological samples include blood (including cord-blood), urine, meconium and faeces, and placental tissue. In subgroups detailed data collection, such as continuous glucose monitoring and videos of parent-child interactions, are being collected.

Keywords

ALSPAC, Birth Cohort, Cross-generation, Data Sharing



This article is included in the [Avon Longitudinal Study of Parents and Children \(ALSPAC\)](#) gateway.

Corresponding author: Deborah A. Lawlor (d.a.lawlor@bristol.ac.uk)

Author roles: **Lawlor DA:** Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Investigation, Methodology, Writing – Original Draft Preparation; **Lewcock M:** Conceptualization, Data Curation, Formal Analysis, Project Administration, Writing – Review & Editing; **Rena-Jones L:** Data Curation, Formal Analysis, Writing – Review & Editing; **Rollings C:** Data Curation, Project Administration, Writing – Review & Editing; **Yip V:** Data Curation, Project Administration, Writing – Review & Editing; **Smith D:** Data Curation, Formal Analysis, Project Administration, Writing – Review & Editing; **Pearson RM:** Data Curation, Investigation, Writing – Review & Editing; **Johnson L:** Data Curation, Investigation, Writing – Review & Editing; **Millard LAC:** Data Curation, Investigation, Writing – Review & Editing; **Patel N:** Data Curation, Investigation, Writing – Review & Editing; **Skinner A:** Data Curation, Investigation, Writing – Review & Editing; **Tilling K:** Conceptualization, Formal Analysis, Writing – Review & Editing;

Competing interests: DAL receives support from several national and international government and charitable funders and from Roche Diagnostics for research unrelated to that presented here. LJ currently receives support from various UK government funding bodies and in the past from Alpro (2015-2016), Kellogg Europe (2015-2016), and Danone Baby nutrition (2009-2012) for research unrelated to that reported here. All other authors report no conflicts of interest.

Grant information: ALSPAC receives core-funding from the University of Bristol, Wellcome Trust and MRC (102215/2/13/2). This study was supported by the National Institute for Health Research NIHR Biomedical Research Centre at the University Hospitals Bristol NHS National Health Service Foundation Trust and the University of Bristol, the US National Institute for Health (grant R01 DK10324), and the European Research Council under the European Union's Seventh Framework Programme (grant FP/2007-2013) / European Research Council ERC Grant Agreements (Grant grants refs 758813 MHINT and 669545 DevelopObese). Medtronic Limited provided support for the pilot study of continuous glucose monitoring. The Wellcome Trust Institutional Strategic Support Fund and Elizabeth Blackwell Institute at the University of Bristol provided support for the novel diet methods pilot study. LACM is funded by a University of Bristol Vice Chancellor's Fellowship. DAL, RP, LACM, AS and KT work in a Unit that is supported by the UK Medical Research Council (MC_UU_00011/6 and MC_UU_00011/3) and University of Bristol, and DAL is a National Institute for Health Research Senior Investigator (NF-SI-0611-10196). The funders had no role in designing the study, collecting or analysing data or contributing to writing the paper. The views expressed in this paper are those of the authors and not necessarily any funding body.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2019 Lawlor DA *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution Licence](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Lawlor DA, Lewcock M, Rena-Jones L *et al.* **The second generation of The Avon Longitudinal Study of Parents and Children (ALSPAC-G2): a cohort profile [version 1; referees: awaiting peer review]** Wellcome Open Research 2019, 4:36 (<https://doi.org/10.12688/wellcomeopenres.15087.1>)

First published: 20 Feb 2019, 4:36 (<https://doi.org/10.12688/wellcomeopenres.15087.1>)

Introduction

Why was the cohort set up?

The Avon Longitudinal Study of Parents and Children-2nd Generation study (ALSPAC-G2) was set up to provide a unique multi-generational family study and to be a resource for international researchers to explore the environmental, socioeconomic, lifestyle, physiological, metabolic, genomic and epigenomic contributions to health and development across the lifecourse and across generations. It builds on the existing ALSPAC resource which originally recruited 14,541 pregnancies to women who were resident in the former county of Avon (centred around the city of Bristol in the South West of England) and who were expected to deliver between 01/04/1991 and 31/12/1992. Those women and their partners (ALSPAC-G0), together with their index children (ALSPAC-G1), who are now in their late-20s, have been followed since pregnancy or birth, with full details provided in previous cohort profiles^{1,2}. The ALSPAC resource, including ALSPAC-G2, receives core funding from the University of Bristol, Wellcome and UK Medical Research Council, with additional support from a very wide range of national and international funders (a comprehensive list of grant funding is available on the ALSPAC website: <http://www.bristol.ac.uk/alspac/external/documents/grant-acknowledgements.pdf>). Since its inception in the early 1990s the study has been known by two names. ALSPAC, including ALSPAC-G2, is used in all academic publications, presentations and with research funders. To participants and in the media the original study (specifically ALSPAC-G1) is known as Children of the 90s (Co90s) and ALSPAC-G2 is known as Children of the Children of the 90s (CoCo90s; sometimes abbreviated verbally to CoCos).

It was envisaged that initially ALSPAC-G2 would contribute unique research in the following broad areas:

1. Understanding how socioeconomic, lifestyle, pathophysiological, metabolic, genomic and epigenomic factors combine to influence the associations of health and wellbeing across generations. With the addition of ALSPAC-G2, ALSPAC is unique in being the only human study able to do this with relevant detailed data across three generations.
2. Determining the impact of pre-conceptual health and wellbeing of mothers and fathers on fertility and growth, development and health of their offspring. The importance of pre-conceptual cohort studies is increasingly recognised but they are difficult to establish, with many existing pre-conceptual cohorts recruited from fertility clinics rather than the general population as here³. For ALSPAC-G2 children, we have detailed repeatedly assessed preconceptual data for at least one of the parents (the original ALSPAC-G1 participant who has become a parent). For the second parent who was not originally in ALSPAC but is recruited to ALSPAC-G1 as part of extending the resource to ALSPAC-G2 we will have pre-conceptual data from record linkage and for subsequent pregnancies/children (i.e. those recruited after the first child that is recruited to ALSPAC-G2) we will also have preconceptual data from our research

clinic. No existing birth cohort has such extensive parental data, including on fathers, with many having little or no data on large proportions of fathers⁴.

3. Providing an opportunity to study people who become parents at a relatively young age and comparing them to couples who become parents at older ages and who are from the same birth cohort and born in the same geographical area.
4. Describing environmental (e.g. climate change, air-pollution), societal (e.g. different methods of communication, changing patterns of between and within country migration), political (e.g. distrust of experts, departures from traditional political parties, changes in gender politics, retirement age and types of employment) technological (e.g. widespread use of information technology and the emergence of artificial intelligence) and lifestyle factors (e.g. sedentary behaviour, reduced smoking and alcohol consumption, increased vaping) that have changed between the current generation of young adults (ALSPAC-G1) and their parents' generation (ALSPAC-G0). Exploring the influence of generational differences in exposure to these factors on decisions about whether, and when, to start a family, parenting patterns, child (ALSPAC-G2) development, parental and child mental and physical health and 'work-family/outside of work' balance.
5. Exploiting the very detailed genetic and phenotypic data collected on parents and siblings to improve causal inference in research, by triangulating results using different methodological approaches, such as Mendelian randomization, parental negative control studies and within sibship analyses⁵.
6. Providing a platform for pilot and feasibility studies of novel technologies for data collection, such as the use of non-intrusive wearables (e.g. biosensors and smart watches) for continually monitoring physiology and behaviours.

Who is in the cohort?

We began recruitment to ALSPAC-G2 on the 6th June 2012. We aim to recruit all children of ALSPAC-G1 participants into ALSPAC-G2. As well as recruiting the next generation of children, we also recruit the (non-ALSPAC) partners of their ALSPAC-G1 parent and collect data from them (Figure 1). Of the 810 ALSPAC-G2 pregnancies/children recruited to date both parents of 74 (9%) were original ALSPAC-G1 children. Thus, for each ALSPAC-G2 pregnancy/child we have very detailed repeatedly assessed information on at least one of their parents from when that parent was *in utero* to the time of them becoming pregnant/a parent (with this being the case for both parents for 9% of ALSPAC-G2). We also have newly collected information from questionnaires, clinic assessments, record linkage and blood samples (on which genome-wide, epigenetic and metabolomic data are being assessed), from both parents, including any parent who was not an original ALSPAC-G1 participant, but is recruited to be part of ALSPAC-G1 as we recruit their ALSPAC-G2 child.

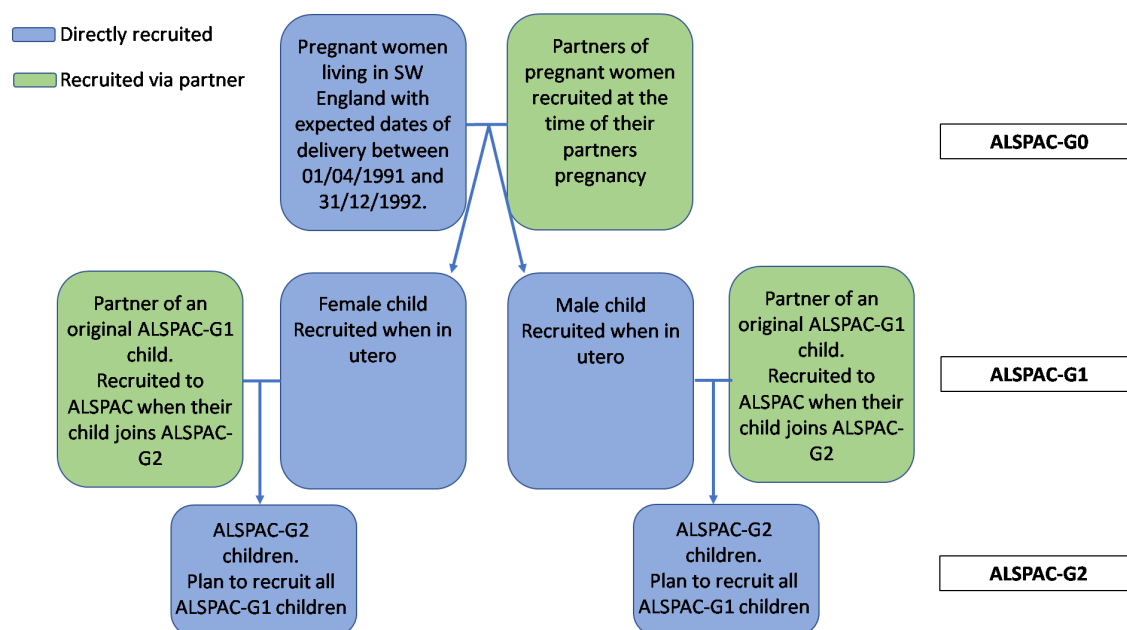


Figure 1. Summarising timing and nature of recruitment to different ALSPAC generations.

We recruit ALSPAC-G2 participants through multiple means: via questionnaires to all ALSPAC-G0 and -G1 participants, in which we ask about becoming pregnant or a parent (ALSPAC-G1) or a grandparent (ALSPAC-G0); asking similar questions when any ALSPAC-G0 or -G1 participant attends a research clinic visit; and through posters in primary care clinics and maternity units in the South West of England. We also use our regular mailed and emailed newsletters and conventional (e.g. local radio and newspapers) and social (e.g. Facebook, Instagram and Twitter) media to remind ALSPAC-G0 and -G1 participants about ALSPAC-G2. When we hear about an ALSPAC-G1 participant becoming pregnant or a parent, from one of their parents (ALSPAC-G0) we ask that their parent mention ALSPAC-G2 to them and ask them to contact us if they are interested in hearing more about the study via an information pack. Similarly, when ALSPAC-G1 participants contact us because of hearing about the study via our newsletter or other media, we ask if they would like an information pack. Parents of ALSPAC-G2 participants can be resident anywhere, though the vast majority remain in the UK; we offer travel costs to clinic visits in Bristol. Biological samples at birth (cord-blood and placental tissue) are currently collected on participants delivering at one of nine maternity units in and around the South West of England.

Ethics. Ethical approval for ALSPAC was obtained from the ALSPAC Ethics and Law Committee and the UK National Health Service Research Ethics Committee (full details are available at: <http://www.bristol.ac.uk/alspac/researchers/data-access/ethics/lrec-approvals/#d.en.164120>). Participants (the main care-giver for children) provided written informed consent for data collection and its use in research.

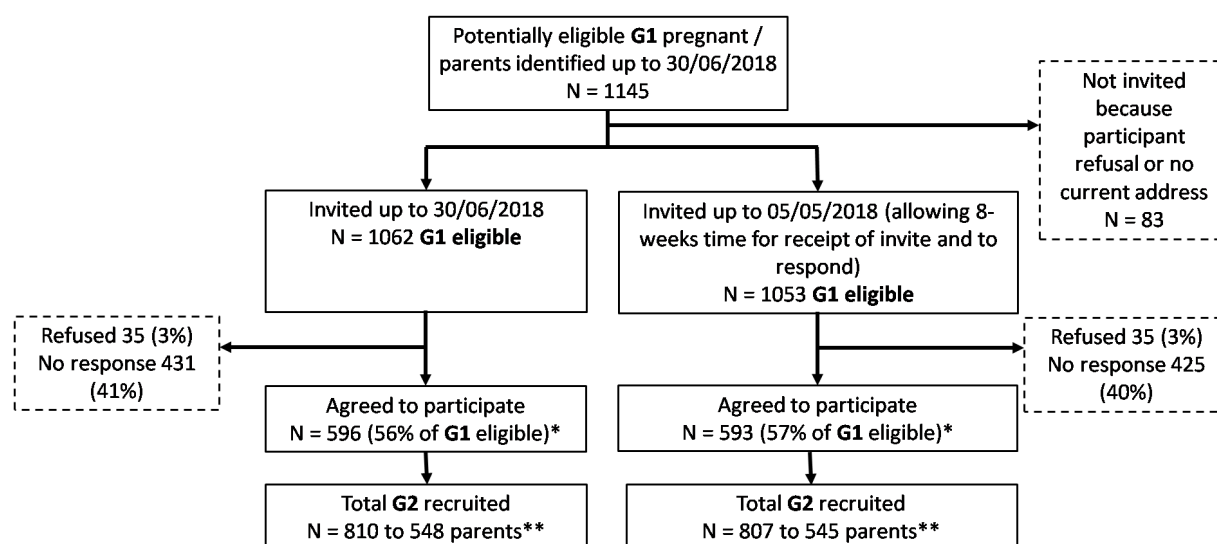
Response and characteristics of eligible participants who are and are not recruited. For all analyses in this profile we have

restricted the sample to children of any ALSPAC-G1 pregnant/parent participant who agreed to be sent an invitation pack and for whom we had a correct address (Figure 2) up to and including 30/06/2018. This provides just over 6 years of data and means we can include all data that has had appropriate quality control checks and linkage to the existing ALSPAC-G0 and -G1 data.

Over this period, we have identified 1145 ALSPAC-G1 participants who were pregnant, had a partner who was pregnant or had become a parent, and we sent invites to 1062 (93%) of these, with 596 (56% of those invited) agreeing to participate. By 30th June 2018 we had assessed and obtained parental consent for 810 ALSPAC-G2 children from 548 families (Figure 2). Of these 810 ALSPAC-G2 participants 389 (48%) were recruited in either early or late pregnancy, with the proportion recruited in pregnancy increasing over time from 20% to 63% between the first year (6th June 2012 to 30th June 2013) to the most recent year (1st July 2017 to 30th June 2018) (Table 1). As might be expected, none of the ALSPAC-G2 participants were recruited at the 7–15-days assessment; all participants seen at 7–15-days were recruited in pregnancy. Similarly, none of the ALSPAC-G2 participants were recruited at the 11-year assessment; those seen at that age were all recruited at earlier ages (Table 1).

Characteristics of those recruited

In analyses using observed data and not taking account of missing data, eligible ALSPAC-G1 parents/pregnant women who were recruited were slightly younger, had higher BMI and were more likely to have attended the two most recent clinic assessments (17–18 or 23–24 years) than those who declined or did not respond, but were similar in terms of sex, educational attainment and smoking (Table 2). However, there was more missing data for these characteristics among those who were not recruited.



* Some of the eligible participants for these periods have subsequently been recruited and our overall response is 60%

** This figure only includes those who have attended their first assessment and consented to be in the study by the given cut-off date; others had assessments booked but had not yet attended

Figure 2. ALSPAC-G2 recruitment from first recruitment (06/06/2012) to 30/06/2018.

We therefore used multivariable multiple imputation to explore the extent to which missing data may have biased these results. Missing values were imputed using chained equations based on ALSPAC-G1 adolescent BMI, and characteristics of their mothers (i.e. ALSPAC-G0; educational attainment, home ownership, occupational social class, parity, smoking during pregnancy, pre-pregnancy body mass index, age at birth of their ALSPAC-G1 child) in addition to all variables in Table 2⁶. For each variable with missing data, 100 imputed variables were created. Linear regression was used to impute continuous variables, logistic for binary, and multinomial logistic for multi-category variables⁶. Differences (differences in means or odds ratios) between characteristics were estimated in each of the 100 imputed datasets and then pooled using Rubin's rules⁶. The results based on observed and imputed data were similar (Table 2).

How often have they been followed-up?

From the start of the study, we established ALSPAC-G2 as an open cohort: open both in terms of when someone might enter the study (i.e. mothers' pregnancy, infancy, childhood or later) and the length of time we aim to keep recruitment open for ALSPAC-G2 (and subsequent generations). We currently have protocols for data collection (on both parents (ALSPAC-G1) and their offspring (ALSPAC-G2)) for early pregnancy (up to 18 completed weeks of gestation), late pregnancy (≥ 28 weeks gestation), first 7–15 days, 6-months, 12 months, annually up to 7 years and then at 9 and 11 years (Figure 3). Our current plans are to extend this with protocols for data collection every 2 years up to the age of 21 over the next 5-years. We also plan to have just one 'pregnancy' assessment at any time during pregnancy (see below).

Participants are invited to all subsequent assessments after the first one that they attend. Both at recruitment and follow-up we are flexible about which assessment age they 'slot' into in order to maximise response and minimise participant burden, particularly when ALSPAC-G1 parents have more than one ALSPAC-G2 child. For example, an ALSPAC-G1 woman who was recruited when 21 weeks pregnant with a second child and who already had a 31-month-old child in ALSPAC-G2, would have the pregnancy assessment for their 2nd pregnancy at the same time as the next assessment due for the first child. That is, we would do the 36-month assessment for the first child a little earlier than had their mother not been pregnant with a second child to minimise participant burden. The family would subsequently be eligible for annual child assessments of the older child and birth, 6-months and then annual child assessments of the younger child. We would plan all subsequent visits so that the family would only need to attend once for each of these subsequent assessments of both children. At all assessments we would complete relevant parental and children assessments and within the study database record this. For example, at the assessments where the mother is pregnant with a second child and also due to have a ~36-month postnatal assessment in relation to their first child, results would be linked to both children and we would do all 36-month measurements (that were feasible in a pregnant woman), as well as all pregnancy measurements. For the 36-month postnatal measurements we would indicate that the woman was pregnant with another child.

We have found that we see many ALSPAC-G1 pregnant women between the thresholds that we initially used to defined early

Table 1. Number of participants recruited at different assessment ages.

Participant recruitment stage	Number of ALSPAC-G2 recruited at each age by year of recruitment (% recruited at each age by each year)						
	Year 1 6 th June 2012–30 th June 2013, n (%)	Year 2 1 st July 2013–30 th June 2014, n (%)	Year 3 1 st July 2014–30 th June 2015, n (%)	Year 4 1 st July 2015–30 th June 2016, n (%)	Year 5 1 st July 2016–30 th June 2017, n (%)	Year 6 1 st July 2017–30 th June 2018, n (%)	Total recruited at each age 6 th June 2012–30 th June 2018
Early pregnancy ^a	8 (6)	30 (25)	21 (23)	51 (33)	53 (32)	68 (44)	231
Late pregnancy ^a	18 (14)	23 (19)	27 (30)	28 (18)	32 (20)	30 (19)	158
<i>Total pregnancy^a</i>	<i>24 (20)</i>	<i>53 (44)</i>	<i>48 (53)</i>	<i>79 (51)</i>	<i>85 (52)</i>	<i>98 (63)</i>	
7–15 days	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0
6 months	30 (24)	19 (16)	17 (19)	35 (23)	18 (11)	18 (12)	137
12 months	26 (21)	18 (15)	5 (6)	7 (5)	15 (9)	8 (5)	79
24 months	18 (14)	13 (11)	10 (11)	13 (9)	7 (4)	10 (6)	71
36 months	14 (11)	9 (8)	3 (3)	5 (3)	13 (8)	8 (5)	52
48 months	6 (5)	3 (3)	4 (4)	7 (5)	13 (8)	0 (0)	33
60 months	4 (3)	4 (3)	2 (2)	2 (1)	4 (2)	3 (2)	19
6 years	3 (2)	0 (0)	0 (0)	2 (1)	2 (1)	5 (3)	12
7 years	0 (0)	1 (1)	1 (1)	3 (2)	6 (4)	2 (1)	13
9 years	0 (0)	0 (0)	0 (0)	0 (0)	1 (1)	4 (3)	5
11 years	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0
Total recruited in each year	127 (100)	120 (100)	90 (100)	153 (100)	164 (100)	156 (100)	810

^a For the period covered in this paper early pregnancy was defined as up to 18-weeks of complete gestation and later pregnancy as 28- or more weeks. However, we have been flexible with these definitions to maximise recruitment and minimise participant burden and found that we see women across all gestations (including between 18- and 28-weeks gestation), thus those included in the early pregnancy category will include women up to 23-weeks and those as late 24- or more weeks. We are currently changing our protocols to having just one assessment during pregnancy at whatever gestational age best suits the woman and her family (see also text in section about how often participants are followed-up). For this reason, we have also provided the total numbers and percentage recruited at any time in pregnancy (in italics).

and late pregnancy (i.e. between 20 and 28 weeks of gestation). This is because once women contact us to join the study they are keen to make an appointment as soon as possible at a date that suits them, and also because of combining assessments when parents have more than one ALSPAC-G2 child (see example description above). We therefore plan to change to having a protocol for recruitment at any gestational age in pregnancy and to see women just once during pregnancy.

In an open cohort like ALSPAC-G2 in which participants may be initially recruited at different ages from when they are *in utero* early in their mother's pregnancy to late childhood, follow-up rates can be difficult to describe. This is because at each of the ALSPAC-G2 assessment ages eligible participants include any fetus/infant/child who is being recruited and seen for the first time at that age and those who were recruited at any of the earlier assessment times who have reached the age of the assessment under consideration. We summarise this in [Figure 4](#). Taking 12 months as an example, we can see that over this period, 496 ALSPAC-G2 participants were eligible for a 12-month assessment

and of those 31% had been recruited in early pregnancy, 26% in late pregnancy, 26% at 6-months and 17% were being recruited for the first time at 12 months. In total 405 had been invited by June 2018 and of those 323 (80%) were assessed. For the first six assessment periods (early pregnancy to 24 months), 71–80% of those invited have attended and attendance remains at 60% or higher for all assessments up to 7-years. For the two oldest ages at which we currently assess ALSPAC-G2 participants, 9 and 11 years, the proportions are 22% and 80%, respectively. The low response at 9 years was because of delays in developing protocols and gaining ethical approval for the later assessment periods, which meant that some had become closer to 11 years and were seen for the first time at that assessment. That said, for both the 9- and 11-year assessments numbers of invited participants are low, and the percent assessed less reliable than at younger ages.

What has been measured?

We use REDCap software for direct data entry at the point of collection. [Figure 3](#) summarises the core data that we currently

Table 2. Comparison of those recruited and those who have declined or not responded. All results are all from questionnaires, research clinic assessments or record linkage data on a ALSPAC-G1 participant who has become pregnant, or for male ALSPAC-G1 has a partner who has become pregnant or who had become a parent and were eligible to be recruited to ALSPAC-G2 up to 30th June 2018.

<i>Continuous outcomes</i>	Recruited ^a (N = 580)		Declined or did not respond ^b (N = 447)		Difference in mean (95%CI) ^c Observed (complete) data	Difference in mean (95%CI) ^c Multivariable imputed data
	N (%) with data	Mean (SD)	N (%) with data	Mean (SD)		
Age (Years) ^d	580 (100)	22.5 (1.9)	446 (100)	22.8 (1.9)	0.37 (0.13, 0.60)	0.37 (0.13, 0.60)
BMI (Kg/m ²) ^e	467 (81)	26.3 (6.1)	215 (48)	25.4 (5.7)	-0.87 (-1.8, 0.1)	-0.91 (-0.05, -1.18)
<i>Binary outcomes</i>	Recruited ^a (N = 580)		Declined or did not respond ^b (N = 447)		Odds ratio (95%CI) ^c Observed (complete) data	Odds ratio (95%CI) ^c Multivariable imputed data
	N (%) with data	Number with the outcome (%)	N (%) with data	Number with the outcome (SD)		
Female	580 (100)	444 (77)	447 (100)	329 (74)	0.85 (0.64, 1.14)	No missing data
Educational attainment to A-level or higher ^f	401 (69)	218 (54)	222 (50)	115 (52)	0.9 (0.65, 1.25)	0.88 (0.63, 1.22)
Ever smoked ^g	509 (88)	336 (66)	252 (56)	170 (67)	1.07 (0.77, 1.47)	1.07 (0.76, 1.49)
Attended 17/18-year or 24/25-year follow-up	568 (98)	472 (83)	441 (99)	221 (50)	0.2 (0.15, 0.27)	0.2 (0.15, 0.27)

^a ALSPAC-G1 participants who were invited and have attended at least one assessment; ^b ALSPAC-G1 participants who were invited and have declined or not responded;

^c Differences in means for age and BMI, odds ratios for educational attainment, smoking and attendance at recent clinics, in all analyses those recruited are the reference category; ^d At time of invitation; ^e At either of the two most recent assessments 24/25 years for those with data from that follow-up, otherwise 17/18 years; ^f A-levels: Advanced-levels, secondary school exams required for University entrance and some other further education/apprentice schemes taken at age 18. Data obtained from record linkage; ^g Data from any previous questionnaire (data on smoking have been collected repeatedly since age 17 years).

collect on parents (ALSPAC-G1) and their offspring (ALSPAC-G2). More detailed lists of variable types are provided in the appendices. These list the information collected via questionnaires, clinical assessments, extraction of data from clinical records and record linkage, and also list the stored biosamples we have for mothers, fathers and the ALSPAC-G2 children (see [Extended Data, Summary of data collection](#))⁷.

We collect a range of environmental, social, lifestyle, clinical and biological data on all family members repeatedly ([Figure 3](#)). Biological samples collected routinely include blood (including cord-blood), urine, meconium and faeces, and placental tissue. Samples are collected with broad generic consent to enable a wide range of future use, including genetic analyses. Up to June 2018, we also invited all relevant mothers (those recruited during pregnancy, 6- and/or 12-months postnatal, and who breast-fed) to provide repeat breast milk samples as part of a pilot study to assess response to this request. In total 457 mothers of 593 children (taking account of siblings and twins) were eligible and of these, 137 mothers (of 168 children) gave consent to

participate in this pilot and 105 (of 121 children) have provided at least one breast milk sample. Of the 105 mothers who provided breast milk, 90 gave sample(s) for 1 infant, 14 gave for 2 infants, and 1 for 3 infants. For any given infant the range of samples that have been obtained is 1 to 4 repeats. To date, no analyses have been performed on these samples (we would be keen to speak to any collaborators who would like to use these samples). We have also collected, or are collecting, intensive repeat or continuously measured data, using novel methods on subgroups, including continuously measured glucose on mothers during pregnancy⁸, videos of parent-child interactions from head-worn and home-based cameras⁹, and dietary intake from smart-phone photographs of meals¹⁰.

Participant engagement and involvement

We have had active participant groups in ALSPAC since it began. These groups suggest areas of data collection and provide advice on planned data collection, including methods for collecting these data. In addition, we receive direct advice and suggestions as we collect data (questionnaires or clinic assessments).

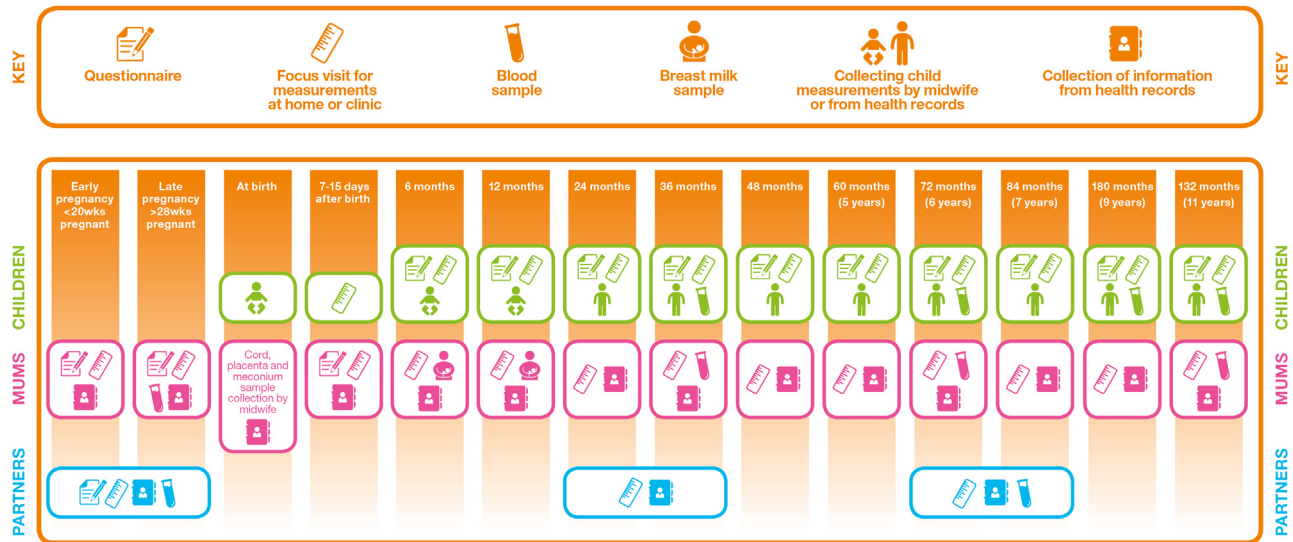


Figure 3. Summary of data collected on parents and children in ALSPAC-G2. This summarises data collection times and type for the period covered by this profile i.e. from initiation of ALSPAC-G2 in June 2012 to the end of June 2018. From Spring 2019 we will be undertaking just one assessment during pregnancy and this will be at a time that best suits the pregnant woman.

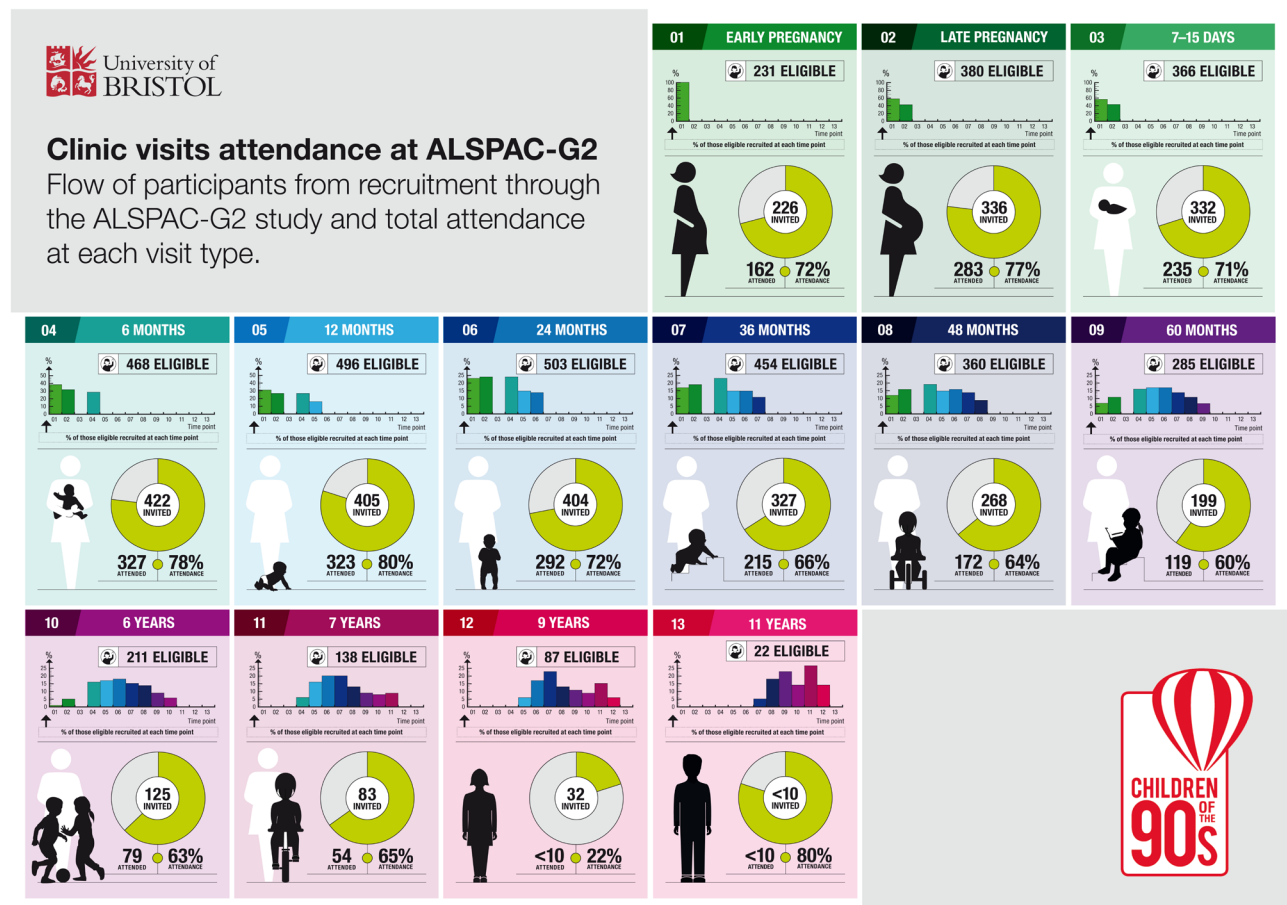


Figure 4. Summary of eligible and invited participants at each assessment time. Each section in this figure represents one of the ALSPAC-G2 age periods (from early pregnancy to 11 years) at which we currently collect data. In each section, for each assessment we report the number eligible and the composition of that eligible number based on when they were first invited and recruited to the study (bar graph at the top of each assessment period). We also show the total number invited to that assessment and the number and percentage of those invited who attended (central figure). All numbers in this figure refer to invites and recruitment between June 2012 and June 2018.

For ALSPAC-G2, this has included help with collecting and appropriately labelling data for same-sex couples and suggesting that we collect data on baby-led weaning and helping us to develop appropriate questions about this. From age 9 years we involve the ALSPAC-G2 children directly via a participant information sheet that they (and their parents) helped us to design (see [Extended Data, Child PIS](#))⁷.

What has ALSPAC-G2 found? Key findings and publications

Cross-generational comparisons of antenatal and infancy characteristics. In pregnant women aged 19–24 years, we have shown that the risk of antenatal depressive symptoms is 50% higher in young women who were pregnant between 2012 and 2016, compared with their mother's generation who were pregnant between 1990 and 1992 (relative risk 1.51 [95% confidence interval: 1.11, 2.05]), with these findings remaining unchanged in numerous sensitivity analyses, including when restricting analyses to 66 mother-offspring pairs¹¹. When individual symptoms were examined, the contemporary generation reported notably higher levels of feeling overwhelmed, crying often and having difficulty sleeping. Offspring of mothers who experienced high levels of antenatal depressive symptoms were over three times more likely to also experience high levels of symptoms¹¹.

We have undertaken additional preliminary cross-generational analyses of pregnancy and perinatal outcomes for this cohort-profile. In these analyses we only included ALSPAC-G0 women who were aged 19–26 at the birth of their ALSPAC-G1, as over 95% of the ALSPAC-G1 women were within this age range when pregnant with their ALSPAC-G2 child. In both groups we only included the first pregnancy recruited to the study. For these preliminary analyses we have only adjusted for maternal age (in years) and we used robust standard errors because of non-independence between the 197 mother-daughter (ALSPAC-G0-ALSPAC-G1) pairs across the two groups. We imputed missing data using the same method as described above for analyses presented in [Table 2](#), except that we used augmented regression approach with small weights to prevent perfect prediction for smoking variables¹².

These preliminary analyses suggest that the current generation of pregnant women (ALSPAC-G1 mothers with their G2 child) are slightly younger, have a higher body mass index and total cholesterol. They also have markedly higher odds of completing education to at least advanced (A)-level secondary school qualifications and markedly lower odds of smoking during pregnancy than their mother's generation ([Table 3](#)). Their (ALSPAC-G2) children are more likely to be delivered by Caesarean section, have a higher mean birth weight and are more likely to have been breast fed as infants than the ALSPAC-G1 generation. Pregnancy haemoglobin levels and gestational age at delivery are similar between the two generations ([Table 3](#)). We do not have pregnancy glucose levels for the ALSPAC-G0 generation but present these for the ALSPAC-G1 generation in [Table 3](#). The results are consistent when based on observed data or multivariable imputed data.

Pilot data collection of physical activity and diet. We piloted the use of a custom made wrist-worn device, which includes a triaxial accelerometer sensor, low-power radio, battery and non-volatile memory module, on 97 mothers¹³. The motivation was to be able to collect very detailed activity data over extended periods of time with as little inconvenience to the participants as possible. The device stores accelerometer data in non-volatile memory allowing for those data to be retrieved (some time later) over a low-power wireless link. It uses a novel, low power, lossless data compression algorithm to minimise the amount of data transmitted over the radio link, whilst retaining all information captured by the accelerometer. This, combined with the low-power radio, makes it possible to collect very detailed activity data whilst minimising overall device power consumption, reducing the burden of having to keep the device charged¹³. However, the pilot study found that participants were unlikely to wear this device because of its large size meaning that they found it less 'attractive' than modern smart watches. We are now collecting pregnancy physical activity levels using the Axivity AX3 wrist-worn tri-axial accelerometer with a considerably greater uptake than the custom made accelerometer (84% versus 50%), resulting in those who provided valid data being greater for the Axivity AX3 accelerometer (254 (79%) of the 322 eligible) compared with the custom-made system (263 (46%) of the 576 eligible).

We are one of the first studies to explore the feasibility of using a smartphone food-photography application to assess dietary intake in a general population of young pregnant women¹⁰. The Remote Food Photography Method (RFPM) collects dietary data using the SmartIntake phone application¹⁴. Pregnant ALSPAC-G1 participants (carrying an ALSPAC-G2 fetus) were asked to record 6 days of eating/drinking occasions with this smartphone application. Real-time monitoring and feedback occurred for the first day. This required them to take a photograph before and after each eating/drinking occasion and provide a brief text description of items that are not visible in photos (e.g. butter inside sandwiches). A total of 182 mothers who were recruited at any assessment point during pregnancy were invited to use RFPM and/or an online food diary or recall to collect dietary data for 6 consecutive days. A greater proportion of women agreed to use the online method compared with RFPM (53% vs 22%). However, of those agreeing to use RFPM, more provided data for 4 days or more than those agreeing to complete the online diary or recall assessments (58% vs 29%). Of those using the RFPM, most found installation and set-up (95%), taking photos of meals (70%) and receiving reminders (81%) easy or very easy¹⁰.

As the use of expert analysis of food photos is extremely research-resource-intensive, we also completed pilot work comparing portion size and food groups identified by expert analysis with crowd-sourced data (using photographs from non-ALSPAC volunteers). We used 30 photographs of meals. For each of these photos total meal weight was measured using a Mandometer® device (Mikrodidakt AB, Lund, Sweden) and food groups displayed were reviewed by an expert dietitian. In comparison to the measured weight, crowds underestimated meal weight

Table 3. Comparison of pregnancy, birth and infancy characteristics between pregnancies occurring 1990 to 1992 (parent generation; ALSPAC-G0/G1) and those occurring 2012–2018 (contemporary generation; ALSPAC-G1/G2).

Continuous outcomes	G0/G1; pregnancies 1991–1992 (N = 5,287)		G1/G2; pregnancies 2012–2018 (N = 494)			
	N (%) with data	Mean (SD)	N (%) with data	Mean (SD)	Age adjusted difference in mean (95%CI)^a Observed (complete) data	Age adjusted difference in mean (95%CI)^a Multivariable imputed data
Age at pregnancy/delivery (Years)	5287 (100)	23.1 (2.6)	494 (100)	21.7 (2.6)	-1.4 (-1.16, -1.64)	No missing data
Pregnancy BMI (Kg/m ²) ^b	2719 (51)	23.9 (4.4)	257 (52)	25.1 (5.7)	1.50 (0.77, 2.23)	1.36 (0.69, 2.03)
Pregnancy cholesterol (mmol/l)	2283 (43)	4.9 (1.5)	133 (27)	6.4 (1.2)	1.6 (1.3, 1.8)	1.5 (1.3, 1.8)
Pregnancy haemoglobin (g/dL)	4303 (81)	12.5 (0.9)	285 (58)	12.6 (1.1)	0.08 (-0.06, 0.21)	0.07 (-0.05, 0.20)
Pregnancy glucose (mmol/l)	NA	NA	132 (27)	5.0 (1.1)	NA	NA
Gestational age (completed weeks) ^c	5287 (100)	39.4 (2.1)	270 (55)	39.38 (2.01)	-0.01 (-0.26, 0.24)	0.01 (-0.24, 0.27)
Birthweight (g) ^e	5217 (99)	3340 (570)	367 (74)	3410 (581)	96 (35, 158)	100 (39, 161)
Binary outcomes	G0/G1; pregnancies 1991–1992 (N = 5,287)		G1/G2; pregnancies 2012–2018 (N = 494)			
	N (%) with data	Number with outcome (%)	N (%) with data	Number with outcome (%)	Age adjusted Odds ratio (95%CI)^a Observed (complete) data	Age adjusted odds ratio (95%CI)^a Multivariable imputed data
Educated to A-level or higher ^c	4410 (83)	827 (19)	321 (65)	175 (54)	7.53 (5.85, 9.7)	7.65 (5.92, 9.88)
Smoked cigarettes in pregnancy ^d	4953 (94)	1948 (39)	210 (43)	33 (16)	0.25 (0.17, 0.36)	0.27 (0.18, 0.39)
Delivered by Caesarean Section ^e	3245 (61)	453 (14)	274 (55)	53 (19.3)	1.66 (1.20, 2.28)	1.67 (1.22, 2.29)
Ever breastfed ^d	3779 (71)	2548 (67)	326 (66)	259 (80)	2.18 (1.65, 2.89)	2.20 (1.67, 2.89)

Results are all from the ALSPAC-G0/ALSPAC-G1 (G0/G1) and ALSPAC-G1/ALSPAC-G2 (G1/G2) mother-offspring pairs. G0 women were recruited in pregnancy between 1990–1992; G1 are the index female offspring from those pregnancies or the female partners of the index male offspring; G2 are their offspring (grandchildren of G0). In both groups only the first pregnancy recruited to the study are included. There was a greater proportion of pregnancies removed from G1/G2 (217 out of 711 (31%)) than G0/G1 (111 out of 5398 (2%)) reflecting the fact that G1/G2 is an open cohort recruiting all children to the original ALSPAC-G1 cohort.

G1/G2 pregnancies occurred between June 2012 and June 2018. For both G0/G1 and G1/G2 analyses are restricted to women who were aged 19–26 years during their pregnancy (the age range for the majority of the G1 women when they were pregnant with G2 offspring).

^a Difference in mean age unadjusted; all other differences in means or odds ratios adjusted for maternal age. In all analyses G0/G1 are the reference category; ^b For both G0 and G1 women weight was abstracted from the first antenatal clinic visit; height was self-reported for G0 and measured in the ALSPAC clinic for G1; ^c A-levels: Advanced-levels, secondary school exams required for University entrance and some other further education/apprentice schemes taken at age 18. Education data was obtained from self-report for G0 and

by an average of 63 g (95% level of agreement -299 to 174 g), whereas the dietician overestimated by 28 g (95% level of agreement -158 to 214 g)¹⁵. In further analyses, we found that compared with expert dietician review, crowds varying in size from 5 to 50 people identified food groups in the photos with high specificity (mean 98%) but modest sensitivity (68%) i.e. crowds almost always identified foods in the photo correctly but some foods in photos were missed out, explaining the average underestimation of total meal weight by crowds (unpublished findings).

Using continuous glucose monitors in unselected prenatal/postnatal women. We have completed a pilot study of the use of continuous glucose monitors (CGM) in healthy women during pregnancy and postnatally. Women were invited to wear a Medtronic iPro2 CGM on their buttock, abdomen or arm, for 6-days, at up to four time points: in early pregnancy (≤ 20 weeks gestation), late pregnancy (> 28 weeks), and at 6 and 12 months postnatally. A total of 63 women provided 96 valid sets of CGM assessment (25% response), with 41 of these women providing

data at a single time point (most commonly early or late pregnancy), 11 on 2 occasions and 9 on 3 occasions (none provided data at all 4 time points)⁸. While wearing the device, participants were asked to measure their capillary blood glucose by finger prick 4 times daily to calibrate the system and to record mealtimes in a meal diary. Feedback from participants suggests that the requirement to repeatedly test capillary glucose with a finger prick was a disincentive to using this system and in Spring 2019 we plan a further study using a different system that does not require calibration using capillary blood from finger pricks.

We used CGMs to record interstitial glucose ‘continuously’, producing a sequence of measurements for each participant (e.g. the interstitial glucose every 5 minutes, over a 24-hour period for a 6-day period). To analyse these data, researchers tend to derive summary variables such as time spent above or below specific levels. To date, a lack of consistency and transparency of precise definitions used for these descriptive characteristics has hindered interpretation, replication and comparison of results across studies¹⁶. We have developed an open-source software package (GLU) for deriving a consistent set of summary variables from CGM data⁸. GLU performs quality control of each CGM sample (e.g. addressing missing data), derives a diverse set of summary variables covering six broad domains, and outputs these measures to the user. We have used GLU with the ALSPAC-G2 pilot data and shown that overall mean glucose levels were very similar (~5 mmol/l) across the four time points, but that this similarity conceals very different patterns of variation, with greater variability during pregnancy (both early and late) than postnatally and more time spent hypoglycaemic during pregnancy than postnatally. Fasting glucose was, on average, higher 12 months postnatally compared with early pregnancy⁸. We also found that, during pregnancy, higher BMI was associated with higher overall mean glucose levels during both the day and night, higher time spent in hyper-glycaemia during the night and shorter post-prandial time to peak glucose⁸.

What are main strengths and weaknesses?

ALSPAC-G2 makes the whole ALSPAC resource a unique inter-generational scientific resource for researchers globally. Recruiting ALSPAC-G2 provides a pre-conception cohort with very detailed information on at least one parent from when they were *in utero*. For the parent who was not a member of the original ALSPAC-G1 cohort, extension of our record linkage to primary and secondary health care, and school-based educational assessments, will provide some pre-conceptual data, and for those with a subsequent child there will be pre-conceptual data on the second and subsequent children. ALSPAC-G2 also has more data on fathers than in most pregnancy/birth cohorts. An additional strength, is that we will use ALSPAC-G2 participants as a control cohort for local and national clinical cohorts, including the on-going national cleft lip and palate cohort, and newly planned Bristol IVF and congenital heart disease cohorts. We have demonstrated the value of ALSPAC-G2 for piloting and testing the feasibility of novel data collection methods and work closely with other international birth cohorts to share best practice, protocols and replicate, and where appropriate pool data.

Currently the number of ALSPAC-G2 participants is relatively small and whilst we have been able to identify large between generation differences, such as the difference in antenatal depression symptoms and smoking during pregnancy, for more modest but potentially important differences it may be some time before these can be precisely estimated. That said, participant numbers are increasing as the G1 participants approach the current mean age for a first pregnancy in the UK (29 years)¹⁷. At the time we began recruitment, ALSPAC-G1 participants were aged 19–21 years and some had already become parents. Thus, whilst the cohort is currently of relatively young parents, we do not have detailed pregnancy data on those who were pregnant before 19-years (though participants are linked to routine health data from which we extract some pregnancy, labour and birth data).

There are a number of sources of potential selection bias¹⁸. Our per cent response is 60% and respondents are more likely to be those already engaged with ALSPAC as indicated by being more likely to have participated in the two most recent clinic assessments for all ALSPAC-G1 participants. Once recruited a high proportion remain actively engaged with follow-up, though the extent to which this will continue as the cohort ages and we attempt to recruit subsequent generations is unknown (see below). Currently, ALSPAC-G2 are children being born to relatively young parents; though the age range of parents will increase with continued follow-up. Existing data, including on parents and grandparents (ALSPAC-G0 and -G1), can be used to explore potential selection bias from any of the selection features described above, and to inform sensitivity analyses to explore the extent of bias that might occur for any given analyses¹⁸. We would encourage scientists using these data to consider this for each analysis they undertake. We try to minimise participant burden and hence maximise follow-up by coordinating data collection for families with more than one child in ALSPAC-G2 and as discussed below we also plan to increasingly pilot and use remote or passive data collection.

The open nature of the study is a strength as it means we will (potentially) capture similar data across the lifecourse from *in utero* to adulthood on all siblings in this generation. However, it adds considerable complexity to the study. For example, for the preliminary cross-generational analyses presented in Table 3 we only included the first pregnancy recruited to either generation. Because ALSPAC-G2 is an open cohort in which we aim to include all children of the ALSPAC-G1 parents, and here we present data from participants recruited over a 6-year period (compared with the original recruitment of pregnancies to ALSPAC-G0 over less than 2-years), we removed more pregnancies from the ALSPAC-G1/G2 cohort than the ALSPAC-G0/G1 cohort (31% versus 2%, respectively). In future analyses we plan to develop and use methods that can analyse all pregnancies as well as use appropriate methods for dealing with missing data; we would encourage collaborators using these data to also consider its multi-level structure (e.g. repeat pregnancies within women across generations and potentially including complex family structures).

We anticipate that most ALSPAC-G2 participants will continue to be born up to the early 2040s (though as some men continue to father children into old age, and we recruit all children, including some adopted and ‘step’-children, some ALSPAC-G2 children may be recruited even later than this). By the 2040s there will also be ALSPAC-G3 and -G4 participants. Continuing to recruit these participants and generations adds considerable value to the existing ALSPAC resource for the whole scientific community. It has the potential to uniquely address important research questions related to how development and adverse health outcomes are transmitted across generations and how we might intervene to maximise population health. However, it may be difficult to continue to obtain the necessary funds for continued recruitment of, and data collection from, multiple generations, and it may also be difficult to retain participant engagement over prolonged periods of time. To address this, we are pursuing more extensive record linkage and novel remote minimally intrusive data capture, such as via smart-watches¹⁹ and sensors for biomarkers⁸, to provide efficient (and accurate) data collection.

ALSPAC-G2 participants have been recruited at different ages and for those recruited after birth, their mother’s pregnancy information and their birth/infancy information is limited to what we can obtain from record linkage. However, the proportion recruited in pregnancy has increased over time. Having an open cohort raises questions about when to complete assays on stored samples—for example, should we wait until we have 500, 1000, or some higher number of cord-blood samples before we assay DNA methylation on them? Currently, we invite enquiries about using the biosamples and will manage these on a case-by-case basis taking account of participant numbers with a given sample, risk of disclosure and assay/processing features such as the extent of ‘batch’ effects. Our aim is to ensure the resource provides the best data for the widest group of scientists.

There are also important considerations about whether we should use older data collection tools, such as those used to assess mental health and social behaviours in previous ALSPAC generations (with the advantage that direct comparisons across generations can be made), or more contemporary tools that might be considered more valid and allow comparisons with other contemporary cohorts. Currently, we have mostly used similar tools to those used in the previous generation, but over the next 12–24 months we will undertake workshops with experts in relevant fields to explore whether there are some measures for which we should be using more contemporary tools.

The fact that this is a cohort in which at least one parent has been actively engaged with the study for over 20-years may mean that they have ‘learnt’ how to respond to some research questions, such as those related to mental health, or complete some assessments, such as cognitive function tests²⁰. A ‘learning effect’ will be unlikely for direct measurements such as weight, height and biosample assays. The ALSPAC-G2 participants are mostly born in the South West of England (as were all of the original ALSPAC-G1 participants) and are mostly of White European origin. This has some advantages for exploring between generational

differences, but we acknowledge replication of findings from ALSPAC-G2 with other more diverse cohorts will be important.

Data availability

The ALSPAC data management plan (<http://www.bristol.ac.uk/alspac/researchers/data-access/documents/alspac-data-management-plan.pdf>) describes in detail the policy regarding data sharing, which is through a system of managed open access. The steps below highlight how to apply for access to the data included in this paper and all other ALSPAC data.

1. Please read the [ALSPAC access policy](#) (PDF, 627kB) which describes the process of accessing the data and samples in detail, and outlines the costs associated with doing so.
2. You may also find it useful to browse the fully searchable [ALSPAC research proposals database](#), which lists all research projects that have been approved since April 2011.
3. Please [submit your research proposal](#) for consideration by the ALSPAC Executive Committee. You will receive a response within 10 working days to advise you whether your proposal has been approved.

If you have any questions about accessing data, please email alspac-data@bristol.ac.uk.

We are very keen for ALSPAC-G2 data to be enhanced and used by external collaborators and are happy for email queries to D.A. Lawlor (d.a.lawlor@bristol.ac.uk) or M. Lewcock (Melanie. Lewcock@bristol.ac.uk).

Please note that the study website contains details of all the data that is available through a fully searchable data dictionary (<http://www.bris.ac.uk/alspac/researchers/data-access/data-dictionary/>).

Extended data

Details of all of the measurements that are currently being collected on ALSPAC-G2 participants and their parents (ALSPAC-G1), including the available stored biosamples, alongside the participant information sheet that we developed with ALSPAC-G2 children for use from age ~9-years, can be accessed on the Open Science Framework. DOI: <https://doi.org/10.17605/OSF.IO/4APU8>⁷.

Author contributions

DAL is the scientific director for ALSPAC-G2, contributed to designing and writing protocols for the study, planned analyses and conducted some analyses for this paper; she wrote the first draft of the paper and coordinated revisions from co-authors.

ML is the project manager of ALSPAC-G2, contributed to writing protocols for the study and manages all aspects of the study, including obtaining ethical approval, recruitment and data collection from participants; she undertook some analyses for this paper and made intellectual comments and edits to drafts of the paper.

L R-J leads quality control checks of ALSPAC-G2 data, undertook some analyses for this paper and made edits to drafts of the paper.

CR contributed to writing protocols, recruitment and collection of data from participants, and made edits to drafts of the paper.

VY manages processing and storage of all biological samples collected on participants and made edits to drafts of the paper.

DS manages data from ALSPAC-G2 and linkage of those data with the earlier generations (ALSPAC-G0 and G1) data; he undertook some analyses for this paper and made intellectual comments and edits to drafts of the paper.

LJ designed and manages the dietary data collection, including the food photography pilot and made intellectual comments and edits to drafts of the paper.

RMP designed and manages data collection on parenting and mental health and made intellectual comments and edits to drafts of the paper.

LACM contributed to the design of the continuous glucose monitoring pilot study and made intellectual comments and edits to drafts of the paper.

NP contributed to the design of the continuous glucose monitoring pilot study and made intellectual comments and edits to drafts of the paper.

AS manages data collection from 'smart' technologies and made intellectual comments and edits to drafts of the paper.

KT manages the statistical analyses for the study, including design aspects that might influence data analysis and interpretation, contributed to analyses presented here and made intellectual comments and edits to drafts of the paper.

Grant information

ALSPAC receives core-funding from the University of Bristol, Wellcome Trust and MRC (102215/2/13/2). This study was

supported by the National Institute for Health Research NIHR Biomedical Research Centre at the University Hospitals Bristol NHS National Health Service Foundation Trust and the University of Bristol, the US National Institute for Health (grant R01 DK10324), and the European Research Council under the European Union's Seventh Framework Programme (grant FP/2007-2013) / European Research Council ERC Grant Agreements (Grant grants refs 758813 MHINT and 669545 DevelopObese). Medtronic Limited provided support for the pilot study of continuous glucose monitoring. The Wellcome Trust Institutional Strategic Support Fund and Elizabeth Blackwell Institute at the University of Bristol provided support for the novel diet methods pilot study. LACM is funded by a University of Bristol Vice Chancellor's Fellowship. DAL, RP, LACM, AS and KT work in a Unit that is supported by the UK Medical Research Council (MC_UU_00011/6 and MC_UU_00011/3) and University of Bristol, and DAL is a National Institute for Health Research Senior Investigator (NF-SI-0611-10196). The funders had no role in designing the study, collecting or analysing data or contributing to writing the paper. The views expressed in this paper are those of the authors and not necessarily any funding body.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgements

We are extremely grateful to all the families who took part in this study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses.

This paper was written on behalf of the ALSPAC executive, and we are grateful to members of the ALSPAC executive, who are Prof Nic Timpson, Ms Lynn Molloy, Dr Kate Northstone, and Dr Susan Ring (all of the University of Bristol) for helpful comments on an earlier draft.

We are also grateful to Ashley Williamson (University of Bristol) who produced [Figure 4](#).

References

- Boyd A, Golding J, Macleod J, *et al.*: Cohort Profile: the 'children of the 90s'--the index offspring of the Avon Longitudinal Study of Parents and Children. *Int J Epidemiol.* 2013; **42**(1): 111–27.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Fraser A, Macdonald-Wallis C, Tilling K, *et al.*: Cohort Profile: The Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *Int J Epidemiol.* 2013; **42**(1): 97–110.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Messerlian C, Williams PL, Ford JB, *et al.*: The Environment and Reproductive Health (EARTH) Study: A Prospective Preconception Cohort. *Hum Reprod Open.* 2018; **2018**(2): pii: hoy001.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Sharp GC, Lawlor DA, Richardson SS: It's the mother!: How assumptions about the causal primacy of maternal effects influence research on the developmental origins of health and disease. *Soc Sci Med.* 2018; **213**: 20–27.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Lawlor DA, Tilling K, Davey Smith G: Triangulation in aetiological epidemiology. *Int J Epidemiol.* 2016; **45**(6): 1866–1886.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Royston P: Multiple imputation of missing values. *Stata J.* 2004; **4**(3): 227–41.
[Publisher Full Text](#)
- Lawlor DA: ALSPAC-G2 cohort. 2019.
<http://www.doi.org/10.17605/OSF.IO/4APU8>
- Millard LAC, Patel N, Tilling K, *et al.*: GLU: A tool for analysing continuously measured glucose in epidemiology. *bioRxiv.* 2018.
[Publisher Full Text](#)
- Lee R, Skinner A, Bornstein MH, *et al.*: Through babies' eyes: Practical and theoretical considerations of using wearable technology to measure parent-infant behaviour from the mothers' and infants' view points. *Infant Behav Dev.* 2017; **47**: 62–71.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

10. England CY, Lawlor DA, Lewcock M, *et al.*: **Use of the Remote Food Photography Method by young pregnant women: a feasibility study.** *Proc Nutr Soc.* 2016; **75**.
[Publisher Full Text](#)
11. Pearson RM, Carnegie RE, Cree C, *et al.*: **Prevalence of prenatal depression symptoms among 2 generations of pregnant women.** *JAMA Network.* 2018; In Press:to be published July 18th.
12. White IR, Daniel R, Royston P: **Avoiding bias due to perfect prediction in multiple imputation of incomplete categorical variables.** *Comput Stat Data Anal.* 2010; **54**(10): 2267–75.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
13. Pope J, Vafeas A, Elsts A, *et al.*: **An Accelerometer Lossless Compression Algorithm and Energy Analysis for IoT Devices.** *IEEE Wireless Communications and Networking Conference Workshops (WCNCW).* 2018; 6.
[Publisher Full Text](#)
14. Martin CK, Nicklas T, Gunturk B, *et al.*: **Measuring food intake with digital photography.** *J Hum Nutr Diet.* 2014; **27** Suppl 1: 72–81.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
15. Johnson L, England CY, Laskowski P, *et al.*: **FoodFinder: developing a rapid low-cost crowdsourcing approach for obtaining data on meal size from meal photos.** *Proc Nutr Soc.* 2016; **75**.
[Publisher Full Text](#)
16. Danne T, Nimri R, Battelino T, *et al.*: **International Consensus on Use of Continuous Glucose Monitoring.** *Diabetes Care.* 2017; **40**(12): 1631–40.
[PubMed Abstract](#) | [Publisher Full Text](#)
17. Haines N, Office of National statistics: **Births by parents' characteristics in England and Wales: 2016.** 2017.
[Reference Source](#)
18. Munafò MR, Tilling K, Taylor AE, *et al.*: **Collider scope: when selection bias can substantially influence observed associations.** *Int J Epidemiol.* 2018; **47**(1): 226–35.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
19. Skinner AL, Stone CJ, Doughty H, *et al.*: **StopWatch: The Preliminary Evaluation of a Smartwatch-Based System for Passive Detection of Cigarette Smoking.** *Nicotine Tob Res.* 2019; **21**(2): 257–261.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
20. Calamia M, Markon K, Tranel D: **Scoring higher the second time around: meta-analyses of practice effects in neuropsychological assessment.** *Clin Neuropsychol.* 2012; **26**(4): 543–70.
[PubMed Abstract](#) | [Publisher Full Text](#)